

ПРИМЕНЕНИЕ МЕТОДОВ КЛАСТЕРНОГО АНАЛИЗА ДЛЯ ДИАГНОСТИКИ МИТОХОНДРИАЛЬНЫХ ЗАБОЛЕВАНИЙ

Е.Я. Гречанина¹, Ю.Б. Гречанина¹, О.В. Васильева¹, А.И.

Поворознюк², А.Е. Филатова²

1Украинский институт клинической генетики Харьковского национального медицинского университета (УИКГ ХНМУ), Харьков, Украина

2 Национальный технический университет «ХПИ», Харьков, Украина

Математическая задача постановки диагноза митохондриального заболевания (МЗ) включает этап построения диагностической модели на основе анализа исходных данных, состоящий в определении информационно значимых признаков этой патологии. Среди них могут быть результаты различных исследований (биохимических, молекулярных и др.), выражающихся в разных условных единицах измерения, что затрудняет возможность сравнения их диагностической значимости для клинициста.

Цель: поиск наиболее информативного математического метода для определения взаимодействий статистически значимых признаков МЗ.

Материалы и методы. При выполнении УИКГ НИР «Всесторонний анализ эпидемиологии и механизмов экспрессии митохондриальных заболеваний в славянских популяциях Восточной Украины» проанализированы результаты молекулярного, биохимического и ультразвукового обследования 186 пациентов с МЗ. Была сформирована исходная база данных, позволяющая выполнять добавление новых признаков и обновление уже существующих без изменения ее структуры.

Результаты. Основным требованием к искомому методу было получение связи одного множества переменных от второго множества и возможность применения метода неполных данных.

В качестве определения связи между множествами переменных использовали метод расположения объектов в многомерном пространстве признаков. Если предположить, что группы признаков связаны между собой, то в многомерном пространстве признаков объекты будут сгруппированы. Поэтому было принято решение использовать методы кластерного анализа (КА) для определения групп объектов, которые позволяют сделать выводы о наличии связей между множествами признаков.

Предложено использовать метод «Слоре» для кластеризации объектов, которые описываются различными подмножествами входного пространства признаков. Основная идея метода состоит в максимизации глобального критерия разбития объектов на кластеры. При этом коэффициент отталкивания регулирует уровень сходства объектов внутри кластера и финальное количество кластеров.

Выводы. В процессе применения метода КА «Слоре» определена его высокая эффективность для предоставления входных данных о пациентах с МЗ, проанализированы клинико-генетические характеристики и обработан результат их взаимодействий. Это дало возможность математического моделирования наиболее распространенных форм МЗ в регионе Восточной Украины.